Exploring the Relation of Vocabulary Size to TOEIC Reading Scores

Yoko Ichige

Abstract

This study investigated how learners' vocabulary size affects their scores of TOEIC reading section. A TOEIC practice test was administered to 40 female university students as the final course test of a TOEIC reading class. The test consisted of a vocabulary part and the three parts of TOEIC reading section (Part 5, 6 and 7). The students' test scores were statistically analyzed to examine how the vocabulary score and the scores of other three parts were correlated. The statistical results were verified qualitatively by analyzing the students' written comments about what was necessary and important to do in order to improve their test scores.

Introduction

In the present study, the relation of test takers' vocabulary size to TOEIC test scores was explored. The purpose of the study was to reveal that in what way and to what extent test takers' vocabulary size is associated with TOEIC scores. By examining the relation of vocabulary size and each part of reading section of TOEIC, it could also become clearer that what abilities are measured in each part, in other words, what constructs are tested by each reading tasks. Thus, the validity of TOEIC reading section was explored by regarding vocabulary size as one of key factors of TOEIC scores.

The recent studies have approached the several issues about TOEIC (Test of English for International Communication) from a variety of points of view. Some studies pointed out that the different types of instruction and learning lead to the difference in the score gains in English proficiency tests. Wilson (2000) reported in the study on TOEIC that the amount that learners were exposed to authentic use of English influenced their listening ability and reading comprehension while the formal instruction at school was associated to the ability in English usage.

From this observation, Wilson pointed out the possibility that the score of reading comprehension "would necessarily be obscured" (p. 22). In the same vein, Ling, Power and Adler (2014) implied, "differences in the instructional and learning foci may be associated with the differential score gain patterns on the four sections of TOEFL iBT practice test" (p. 14). Thomas et.al. investigated the relationship among learning styles and achievement in grades and TOEIC scores of Japanese learners who were in the oversea program in New Zealand (2000). Their study confirmed the results of other related studies showing "[k]inesthetic, auditory and tactile styles are the top three preferences and group [work] is the lowest" (p. 3). While the relationship between TOEIC scores and learning styles was limited, the positive relationship between learning styles and grades was observed. They explained about the latter result by pointing out the type of assessment used in the course and the grades for the course were based on the students' performance but not on language-based. They also listed the typical preferences of the students, such as "a very strong preference for concrete sequential styles over intuitive …. a preference for global (field-dependent) rather than analytical (field-independent) styles of learning" (p. 4). They assumed that both these preferences indicate "the Japanese education system" and "the importance of the group in Japanese culture" (p. 4). Another study was conducted in Taiwanese technical university on the relation of TOEIC as a criterion for graduation to learning styles between two groups of students (Pan, 2014). One of the two groups was required to take TOEIC for their graduation while the other was not. In general, there was not a significant difference in TOEIC scores and the students' learning styles between the two groups. Both groups preferred traditional learning activities which are rather receptive, such as, studying vocabulary, grammar, reading and listening to English to more productive communication-oriented activities. However, a significant difference was detected between groups devided based on proficiency regarding their learning styles. The higher proficiency group was engaged more in communication-oriented activities than in tests-specific preparation while the lower proficiency group did the opposite. Pan suggested that TOEIC is to measure everyday communication ability; the learners are to be provided with more communicative activities.

The present study approached TOEIC issues from a more psychological construct-specific perspective than those mentioned above. As written earlier, the study regarded test takers' vocabulary size as one of the major factors of TOEIC scores, especially its reading section scores.

Method

Participants

The data were collected from 40 female university students including 6 foreign students from Korea and China at a Japanese women's university located in Tokyo. Most of them were in their second year except for a few third-year students. Their majors were all related to science about human life, such as, Food science, Environment study. Even though English is not required in their majors, and most of the students in the course were not confident in English, they thought it was necessary to learn English, especially to take TOEIC, for their future career. The level of their English was low- intermediate on average. All of them took the general English course required in their first year. As for the students form Korea and China in the course, the level of their English was ranging from intermediate to high-intermediate, that is, much higher than the Japanese students in the course, and had some experience of taking TOEIC and the knowledge about TOEIC.

Course

The course was an elective one, designed for the students to get familiar with TOEIC reading section and develop their skills specific to taking TOEIC as well as their overall communicative English skills. Lessons were conducted based on the course textbook that consists of the exercises of each part of TOEIC reading section. Additional communicative tasks, such as peer-interviews and role plays, were given to the students between the exercises so that they could develop the skills to use English in more realistic way than by working only on the written exercises in the textbook.

Two TOEIC practice tests (about 60 minute-long reading tests) were administered in the middle and at the end of the course. The students were also required to record and report about the reading materials that they read outside the classroom during the semester.

Data collection

The mid-term test and the final test were constructed based on the reading section of TOEIC and administered to the students. Both tests were of the same format composed of four parts, a vocabulary quiz and each part of TOEIC reading section (Part 5, 6 and 7). Each test began with the vocabulary quiz which asked the students to choose the word meanings of those words learned in the lesson (20 words in the mid-term test and 25 words in the final test), followed by Incomplete Sentences, Text Completion and Reading Comprehension of TOEIC reading section.

Besides these tests, the students were asked to write what they thought might affect their scores of TOEIC and how they thought their scores of TOEIC would be like in the future at the end of the course.

Analyses

The scores of the mid-term test and the course final test were collected as the data, which was classified into four kinds of sub data, that is, Vocabulary score, Part 5 score, Part 6 score and Part 7 score, which were analyzed as four interval variables.

Vocabulary score, Part 5 score, Part 6 score and Part 7 score were tested for correlation respectively for the mid-term test total score and the final test total score. At the same time, the correlation among the scores of vocabulary, Part 5, 6, and 7 were also tested for both the mid-test and final test. Then, further analyses were conducted qualitatively to understand in more depth the relationship between vocabulary size and each part of both the mid-term and final test by probing the characteristics of each part, especially of the one that seemed problematic.

Results

As for the mid-term test, all kinds of scores, that is, vocabulary scores, Part 5, 6, and 7 scores were significantly ($p.< .005$) correlated. (Table 2) Also, the tests conducted among the scores showed significant correlations, except between vocabulary score and Part 6 score. Similarly, the correlation tests for the final test resulted in significant correlations among the final test total score, vocabulary score, Part 5, 6, 7 scores, except between vocabulary score and Part 7 score. (Table 3)

Table 1 *Descriptive statistics of mid-term test scores and final test scores*

Mid-tern test				Final test			
	N	*Mean*	*S.D.*		*N*	*Mean*	*S.D.*
Total score	40	60.63	15.48	Total score	40	65.20	15.16
Vocabulary	40	16.60	5.77	Vocabulary	40	11.83	3.49
Part 5	40	8.78	4.45	Part 5	40	9.08	4.58
Part 6	40	11.48	3.46	Part 6	40	22.50	6.11
Part 7	40	14.58	4.83	Part 7	40	17.13	5.62

Table 2 *Correlations among vocabulary, Part 5, 6, 7 and mid-tern test total scores (p. <.005)*

	Part 5	Part 6	Part 7	Mid-term test
Vocabulary	.62*	.27	.39*	.81*
Part 5		.41*	.43*	.82*
Part 6			.45	.65*
Part 7				.65*

Table 3 *Correlations among vocabulary, Part 5, 6, 7 and final test total scores (p. <.005)*

	Part 5	Part 6	Part 7	Final test
Vocabulary	.47*	.44*	.28	.72*
Part 5		.70*	.37*	.77*
Part 6			.56*	.84*
Part 7				.76*

Form the significant correlation among all the scores of the mid-term test and final test, it is safely said that both the mid-term and final tests conducted in the present course measured one construct consistently enough to infer the learner's communicative ability. However, it was necessary to investigate the insignificant correlation between vocabulary score and Part 6 score of the mid-term test as well as between vocabulary score and Part 7 of the final test. Qualitative inquiry was carried out on the items of Part 6 and Part 7 of both the mid-term and final test to understand more precisely about the constructs measured in these two parts, and further to detect the reason for the insignificant correlation with vocabulary score. The analyses revealed that the items of Part 6 of the mid-term test mostly required learners'

knowledge of usage (*e.g.*, the usage of preposition, such as, "*between* the 25th and 30th", "*by* the 15th of this month"), while vocabulary knowledge seemed necessary to answer Part 6 of the final test. Thus, even though the format of Part 6 was same in both tests, the necessary constructs to solve the items were different. Similarly, the items of Part 7 required different constructs in the two tests. Part 7 items of the mid-term test needed vocabulary knowledge to answer as the learners had to know the synonyms of the words questioned, that is knowledge of paraphrasing, such as, "*neighboring*" in the passage and "*next to*" in the answer options. On the other hand, Part 7 items of the final test could mostly be answered based on the understanding the context or organization of the passage.

Discussion

From the results of the analysis, two points were confirmed. Firstly, the mid-term test and final test used in the present study were adequately designed to measure one consistent construct of reading. And the learners with bigger vocabulary size more likely to mark higher scores on these tests. Thus, it was confirmed that vocabulary size could affect the score of TOEIC reading section. Furthermore, the very high correlation between vocabulary score and the total score in both cases of the mid-term test and the final test (.81 for the mid-term test and .72 for the final test) may indicate the possibility that these tests more likely to demonstrate learners' vocabulary size rather than their reading ability.

This inference seems to be associated with what Wilson found if vocabulary is considered as a kind of knowledge. Wilson wrote (2000) that the students with more exposure to English use, for instance, by studying abroad, marked higher scores in the reading section of the TOEIC practice test than those who had more usage-centered formal English education, who marked higher scores in the usage section. Considering the fact that most students of the present study had limited experience of being exposed to real English use, in other words, with little experience of reading English for real purposes, the students may tend to rely more on their knowledge acquired through studying the textbook in the lesson than on the skills acquired through their own reading experiences, such as, trying to infer the meaning of unknown words based on the context or the overall organization of the passage, or using English to get necessary information. Thus, it is not denied that the tests used in the present study could be measuring the students' knowledge more than the ability to read English

passages. So, the finding of the present study seems to support Wilson's finding and may imply that less exposure to real English use may lead the learners to rely more on the knowledge of English usage than on the ability of English use.

This claim was verified by the qualitative data collected at the end of the course as the students' comments on TOEIC reading section. The students were divided into three score groups, the higher-score group, the medium-score group, and the lower-score group. The students wrote what they thought were main factors that might influence their score of TOEIC. "Vocabulary", "grammar", and "experiences of taking tests" were listed by all three groups of students as the main factors affecting TOEIC reading section scores. The students seem to have the subjective feeling that TOEIC reading section requires knowledge about English usage. However, the differences in their comments were found among the score groups. The higher-score group gave "preparation" as one of important factors, which was not referred to by the other two groups. In addition, the higher-score group came up with a variety of factors while the other two groups wrote a limited range of factors. "Technique", "carefulness", "memorizing", "spelling", or "writing" were among the list by the higher-score group. These factors are not knowledge–related but more technical and specific to taking tests. On the other hand, the lower-score group wrote only two factors that were unique to them, "School education until high school" and "teacher". These factors are more or less associated with formal education. Five of the six foreign students from China and Korea in the course were involved in the higher-score group, who were using English as the medium for communication while living in Japan. Therefore, the students of the higher-score group may have had more exposure to real English use than those of the medium-score group and the lower-score groups. These differences in the students' subjective comments also appeared to be associated with Wilson's assertion, that is, the amount of exposure to real English use could lead to learners' different views about their approach to TOEIC reading section.

References

Ling, G., Powers, D. E., & Adler, R. M. (2014). Do TOEFL iBT® scores reflect improvement in English-Language proficiency? Extending the TOEFL iBT validity argument. *Research Report, 14-09*. Princeton, NJ: ETS.

Pan, Y. (2014). Learner washback variability in standardized exit tests. *The Electronic Journal for English as a Second Language, 18*.

Thomas, H., Cox, R. & Kojima, T. (2000). Relating preferred learning style to student achievement. *Paper presented at the Annual Meeting of the Teachers of English to Speakers of Other Languages (Vancouver, BC)*

Wilson, K. M. (2000). AN exploratory dimensionality assessment of the TOEIC test. *Research Report 00-14*. Princeton, NJ: ETS.